

General reference decoder for image of video frequency process

Publication number: CN1426235 (A)

Publication date: 2003-06-25

Inventor(s): ZHOU P A [US]; COLBOLA J R [US]

Applicant(s): MICROSOFT CORP [US]

Classification:






- **International:** *H04N7/26; G06T1/00; H04L13/08; H04L29/08; H04N7/50; H04N7/26; G06T1/00; H04L13/08; H04L29/08; H04N7/50; (IPC1-7): H04N7/24; G06T9/00; H04N11/04*

- **European:** *H04N7/26A6E4E; H04N7/26A4E; H04N7/26A6E2; H04N7/26A8P; H04N7/26D; H04N7/50E5*

Application number: CN20021043213 20020919

Priority number(s): US20010955731 20010919

Also published as:

 CN100461858 (C)
 EP1298938 (A2)
 EP1298938 (A3)
 US2003053416 (A1)
 US2006198446 (A1)

[more >>](#)

Abstract not available for CN 1426235 (A)

Abstract of corresponding document: **EP 1298938 (A2)**

A method and system including an improved generalized reference decoder that operates according to any number of sets of rate and buffer parameters for a given bit stream. Each set characterizes a leaky bucket model and contains three parameters representing the transmission bit rate, buffer size, and initial decoder buffer fullness. An encoder provides at least two sets of these parameters, whereby the decoder selects one or interpolates between them to operate at any desired peak bit rate, buffer size or delay. The generalized reference decoder may select the smallest buffer size and corresponding delay that decodes the bit stream without buffer underflow or overflow, or alternatively may select and operate at the minimum required peak transmission rate, or something between both.; In practice, the buffer size, delay and/or the peak transmission rate can be reduced by significant factors, and/or the signal-to-noise ratio (SNR) can be increased.

Data supplied from the **esp@cenet** database — Worldwide



[12] 发明专利申请公开说明书

[21] 申请号 02143213.9

[43] 公开日 2003 年 6 月 25 日

[11] 公开号 CN 1426235A

[22] 申请日 2002.9.19 [21] 申请号 02143213.9

[30] 优先权

[32] 2001. 9. 19 [33] US [31] 09/955,731

[71] 申请人 微软公司

地址 美国华盛顿州

[72] 发明人 P·A·周 J·R·科尔伯拉

[74] 专利代理机构 上海专利商标事务所

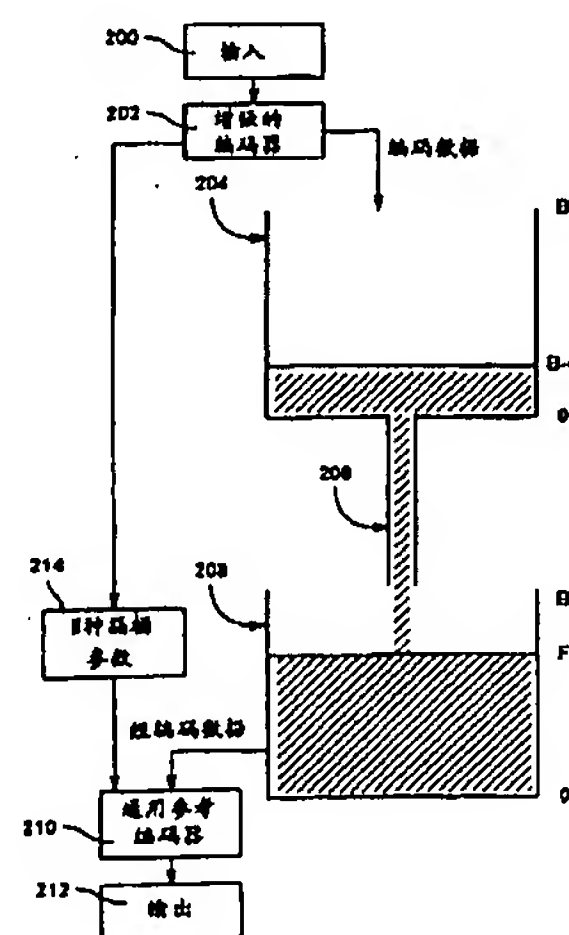
代理人 钱慰民

权利要求书 4 页 说明书 12 页 附图 5 页

[54] 发明名称 用于图像或视频处理的通用参考解
码器

[57] 摘要

一种方法和系统包括一个改进过的通用参考解
码器，该通用参考解码器根据给定位流的任何数量的
各组速率与缓冲器参数来进行操作。 每组参数表
现了漏桶模型的特征，并包含代表传输位速率、缓
冲器大小和初始的解码器缓冲器充满度的三种参
数。 编码器提供至少两组这些参数，由此，解码器
选择一种参数或在它们之间内插，以便以任何所需
的峰值位速率、缓冲器大小或延迟进行操作。 通用
参考解码器可以选择对位流解码而不发生缓冲器下
溢或溢出的最小的缓冲器大小与对应的延迟，或
者，可以按所要求的最小峰值传输速率来进行选择
和操作，或介于两者之间。 在实际应用中，缓冲器
大小、延迟和/或峰值传输速率可以减小很多倍，并
且/或者，可以提高信号—噪声比(SNR)。



1. 一种计算机实现的方法，其特征在于，包括：

使用包括速率数据和缓冲器大小数据的至少两组参数，以便通过选择以下内容来确定一个操作条件，内容为：

1) 基于速率数据的缓冲器大小；或

2) 基于缓冲器大小的速率；

以及，

在随时间变化信号的解码器处，根据该操作条件来保持缓冲器中的编码数据，并对来自缓冲器的编码数据进行解码。

2. 如权利要求 1 所述的方法，其特征在于，还包括：

在随时间变化信号解码器处接收所述至少两组参数，其中，所述随时间变化信号解码器确定操作条件。

3. 如权利要求 2 所述的方法，其特征在于：每组参数也包括在随时间变化信号解码器处所接收的充满度数据。

4. 如权利要求 2 所述的方法，其特征在于：所述至少两组参数由编码器确定。

5. 如权利要求 4 所述的方法，其特征在于：在流头部中接收所述至少两组参数以及指示各组参数的总数的信息。

6. 如权利要求 1 所述的方法，其特征在于：使用所述至少两组参数来确定操作条件，包括选择一组参数。

7. 如权利要求 1 所述的方法，其特征在于：使用所述至少两组参数来确定操作条件，包括在至少两组参数中的各个数据点之间内插。

8. 如权利要求 1 所述的方法，其特征在于：使用所述至少两组参数来确定操作条件，包括在至少两组参数中的各个数据点外插。

9. 如权利要求 1 所述的方法，其特征在于：根据速率数据来选择缓冲器大小，包括确定将要接近最小的加载延迟的缓冲器大小。

10. 如权利要求 1 所述的方法，其特征在于：根据缓冲器大小来选择速率，包括根据缓冲器大小来确定所要求的最小的峰值传输速率。

11. 如权利要求 1 所述的方法，其特征在于：操作条件在经编码数据传送给缓冲器期间至少改变一次。

12. 一种计算机实现方法, 其特征在于, 包括:

在随时间变化信号的解码器处, 接收包括速率和缓冲器大小数据的至少两组参数;

使用所述至少两组参数, 以便通过选择以下内容来确定一个操作条件, 内容为:

1) 基于速率数据的缓冲器大小; 或

2) 基于缓冲器大小的速率;

根据该操作条件来保持缓冲器中的经编码数据; 以及,

对来自缓冲器的经编码数据进行解码。

13. 如权利要求 12 所述的方法, 其特征在于, 还包括: 将充满度数据提供给所述随时间变化信号解码器。

14. 如权利要求 12 所述的方法, 其特征在于, 还包括: 确定至少两组参数并将它们提供给随时间变化信号的解码器。

15. 如权利要求 14 所述的方法, 其特征在于: 所述至少两组参数由编码器确定。

16. 如权利要求 15 所述的方法, 其特征在于: 在流头部中接收所述至少两组参数以及指示各组参数的总数的信息。

17. 如权利要求 12 所述的方法, 其特征在于: 其中, 使用所述至少两组参数来确定操作条件, 包括选择各组参数中的一组参数。

18. 如权利要求 12 所述的方法, 其特征在于: 其中, 使用所述至少两组参数来确定操作条件, 包括在至少两组参数中的各个数据点之间内插。

19. 如权利要求 12 所述的方法, 其特征在于: 其中, 使用所述至少两组参数来确定操作条件, 包括从至少两组参数中的各个数据点外插。

20. 如权利要求 12 所述的方法, 其特征在于: 根据速率数据来选择缓冲器大小, 包括确定将要接近最小延迟的缓冲器大小。

21. 如权利要求 12 所述的方法, 其特征在于: 根据缓冲器大小来选择速率, 包括根据缓冲器大小来确定所要求的最小的峰值传输速率。

22. 如权利要求 12 所述的方法, 其特征在于: 所述操作条件在经编码数据传送给缓冲器期间至少改变一次。

23. 用于提供随时间变化信号的一种系统, 其特征在于, 包括:

提供随时间变化信号的编码器;

保持随时间变化信号的编码器缓冲器和解码器缓冲器，所述编码器缓冲器通过传输媒体连接到所述解码器缓冲器；

将随时间变化信号从解码器中除去的解码器；以及，

确定包括速率数据和缓冲器大小数据的至少两组参数的第一种机制，所述参数组用于保持解码器缓冲器，以便其不会发生溢出或下溢，以及，

第二种机制，根据速率数据来确定解码器缓冲器的大小，或根据缓冲器大小数据来确定数据从编码器缓冲器转传送到解码器缓冲器的速率。

24. 如权利要求 23 所述的系统，其特征在于：其中，确定所述至少两组参数的第一种机制包括在编码器中。

25. 如权利要求 23 所述的系统，其特征在于：其中，所述第二种机制包括在解码器中。

26. 如权利要求 23 所述的系统，其特征在于：其中，确定所述至少两组参数的第一种机制包括在编码器中，第二种机制包括在解码器中；其中，所述编码器将各组参数传送给解码器。

27. 如权利要求 26 所述的系统，其特征在于：其中，所述编码器经由流头部将各组参数传送给解码器。

28. 如权利要求 26 所述的系统，其特征在于：所述编码器识别各组参数的总数。

29. 如权利要求 23 所述的系统，其特征在于：其中，每组参数也包括在解码器处所接收的充满度数据。

30. 如权利要求 23 所述的系统，其特征在于：第二种机制根据速率数据来确定解码器缓冲器的大小，或通过选择一组参数，确定传送数据的速率。

31. 如权利要求 23 所述的系统，其特征在于：第二种机制根据速率数据来确定解码器缓冲器的大小，或通过所述至少两组参数中的各个数据点之间内插，来确定传送数据的速率。

32. 如权利要求 23 所述的系统，其特征在于：第二种机制根据速率数据来确定解码器缓冲器的大小，或通过从所述至少两组参数中的数据点外插，来确定传送数据的速率。

33. 如权利要求 23 所述的系统，其特征在于：其中，通过确定将要接近最小的加载延迟的缓冲器大小，第二种机制可确定解码器缓冲器的大小。

34. 如权利要求 23 所述的系统，其特征在于：通过确定对应于预定缓冲

器大小的、所要求的最小峰值传输速率，第二种机制可确定速率数据。

35. 如权利要求 23 所述的系统，其特征在于：第二种机制根据速率数据和时间信息来确定解码器缓冲器的新大小。

36. 如权利要求 23 所述的系统，其特征在于：第二种机制根据缓冲器大小数据大小和时间信息来确定传送数据的新速率。

用于图像或视频处理的通用参考解码器

发明背景

本发明涉及图像和视频信号以及其他随时间变化的信号（例如语音和音频）信号的解码。

发明背景

在视频编码标准中，至少从概念上讲，如果位流可以通过连接到编码器输出的解码器的数学模型进行解码，那么，它是适应的。这种模型解码器在 H. 263 编码标准中被称作“假设参考解码器”（HRD），而在 MPEG 编码标准中被称作“视频缓冲检验符”（VBV）。一般而言，真实的解码器设备（或终端）包括一个解码器缓冲器、一个解码器和一个显示单元。如果真实的解码器设备根据解码器的数学模型来构制，并且适应的位流在特殊条件下被传输到该设备，那么，解码器缓冲器将不会溢出或下溢，并且解码将得到正确的执行。

前面的参考（模型）解码器假设：位流将以给定的恒定位速率通过信道传输，并且将由具有某给定的缓冲器大小的设备来进行解码（在给定的缓冲延迟之后）。所以，这些模型非常固定，不能达到当今许多重要的视频应用的要求（例如，实况转播视频，或在具有各种峰值位速率的网络路径上提出需要时就将预先编码过的视频发送到具有各种缓冲器大小的设备。

在以前的参考解码器中，按给定的恒定位速率（通常是用位/秒来表达的流平均速率）接收视频位流，并在缓冲器达到某个所需的充满度水平之前存储在解码器缓冲器中。例如，在解码可以从其中重构一个输出帧之前，至少需要对应于视频信息的一个初始帧的数据。这个所需理想的水平表示作为初始的解码器缓冲器充满度，并且在恒定的位速率时，直接与传输或启动（缓冲器）延迟成正比。一旦达到这个充满度，解码器就立即（本质上）除去该序列的第一个视频帧的各个位，并且为这些位解码，以显示该帧。后面各个帧的位也立即按随后的时间间隔除去、解码和显示。

这种参考解码器按固定的位速率、缓冲器大小和初始延迟进行工作。但是，在许多当代的视频应用（例如，通过互联网或 ATM 网络的视频流）中，

峰值带宽根据网络路径而变化。例如，峰值带宽根据与网络的连接是否通过调制解调器、ISDN、DSL、电缆等而有所不同。而且，峰值带宽也可能根据网络条件（例如，根据网络拥塞、所连接的用户的数量和其他已知的因素）而发生及时的变动。此外，视频位流被传送到具有不同的缓冲器容量的各种设备（包括手机、“个人数字助手”（PDAs）、PCs、可放进口袋的计算设备、电视机顶盒、类似 DVD 的播放器等设备），并且为具有不同的延迟要求的情况（例如，低延迟流、渐进下载等类似的情况）而创建。

现存的参考解码器不能对这些变数进行调整。同时，编码器通常事先不知道，也无法知道对于给定的接受者有什么可变的条件。结果，资源和/或延迟时间经常被不必要地浪费，或者在许多情况下并不合适。

发明内容

简而言之，本发明提供了一种改进的通用参考解码器，该通用参考解码器根据给定位流的任何数量的各组速率与缓冲器参数来进行操作。每组参数表现被称作“漏桶模型”（leaky bucket model）或“参数组”的特征，并且包含三个值（R、B、F）——其中，R 是传输位速率，B 是缓冲器大小，F 是初始的解码器缓冲器充满度。不言而喻， F/R 是启动或初始缓冲器延迟。

编码器创建由某所需数量 N 个漏桶所包含的一个视频位流，或者，在已生成该位流之后，编码器可以只计算 N 组参数。编码器将该数量和对应数量的（R、B、F）组以某种方式（例如，在初始流头部中或频带以外）传送给解码器（至少）一次。

当在解码器处被接收时，如果至少有两组参数，则通用参考解码器选择一组参数或在各个漏桶参数之间内插，由此可以按任何所需的峰值位速率、缓冲器大小或延迟进行操作。尤其是，在所给出的所需峰值传输速率 R' （其在解码器末端已知）的情况下，通用参考解码器（根据可用的（R、B、F）组，通过选择一组参数，或在两组参数或多组参数之间内插，或通过外插）选择最小的缓冲器大小和延迟，这种选择将能够对位流解码，而不会遭受缓冲器下溢或溢出的困扰。或者，关于给定的解码器缓冲器大小 B' ，假设解码器将选择和最小的所需峰值传输速率进行操作。

通用参考解码器的好处包括：内容供应商可以创建一个位流一次，而服务器可以使用不同峰值传输速率的各种信道来将它传送到容量不同的多种设

备。或者，服务器和终端可以商议所给定的联网条件的最佳漏桶参数（例如，将产生最低的启动（缓冲器）延迟的参数，或对设备的给定缓冲器大小要求最低的峰值传输速率的参数）。在实际应用中，一些终端的缓冲器大小和延迟可以减小一个数量级，或者，峰值传输速率可以减小一个大的倍数（例如，4 倍），并且/或者，信号-噪声比（SNR）也许可以提高几个 dB，而不会提高平均的位速率（除了传达漏桶信息的可忽略数量的额外的位以外）。

通过以下详细的描述并结合附图，将明白其他的好处和优点。在附图中：

附图说明

图 1 是方框图，表现了可以包含本发明的一种示范的计算机系统；

图 2 是方框图，表现了增强的编码器和通用参考解码器及其它们各自的缓冲器，用于根据本发明的一个方面来为视频或图像数据进行编码和解码；

图 3 是当被包含在一个漏桶的参数（R，B，F）中时缓冲器充满度随时间变化的图表；

图 4 表现了典型视频剪切片断的速率与缓冲器大小的曲线；

图 5 表现了根据本发明的一个方面的典型视频剪切片断的速率与缓冲器大小的曲线，将两个漏桶模型（参数组）提供给通用参考解码器，用于内插和外插。

具体实施方式

示范的操作环境

图 1 展示了可以在其中实现本发明的合适的操作环境 120 的一个例子，特别适用于对图像和/或视频数据进行解码。操作环境 120 仅是一个合适的操作环境的实例，不是对本发明的使用或功能性的范围方面提出任何限制。其他众所周知的、可能适用于本发明的计算系统、环境和/或配置包括（但不局限于）个人计算机、服务器、计算机、手持设备或便携式设备、多处理器系统、基于微处理器的系统、可编程的消费电子产品、网络 PCs、微型计算机、主机计算机、包括以上任何系统或设备的分布式计算环境和类似物。例如，有可能经常在具有超越当代手持个人计算机处理能力的计算机上执行对图像和/或视频图像数据编码，但是，没有理由断定无法在示范设备上执行编码或在功能更加强大的机器上执行解码。

可以在由一台或多台计算机或其他的设备所执行的计算机可执行指令（例如，程序模块）的一般环境中描述本发明。一般而言，程序模块包括实行特定任务或实现特定的抽象数据类型的例程序、程序、对象、组成部分、数据结构等。通常，程序模块的功能性可以按各个实施例的需要来加以结合或分配。

计算设备 120 通常至少包括某种形式的计算机可读媒体。计算机可读媒体可以是可由计算设备 120 进行存取的任何可用的媒体。举例来讲（但不作限制），计算机可读媒体可以包括计算机存储媒体和通信媒体。计算机存储媒体包括易失性与非易失性媒体，以及可移动与不可移动媒体，这些介质媒体是采用可存储信息（例如，计算机可读指令、数据结构、程序模块或其他数据）的任何方法或技术来实现。计算机媒体包括（但不局限于）RAM、ROM、EEPROM、快闪存储器或其他存储器技术、CD-ROM、数字通用光盘（DVD）或其他光学存储器、磁性盒式磁带、磁带、磁盘存储器或其他磁性存储设备，或任何可用来存储所需的信息并可由计算设备 120 进行存取的任何其他的媒体。通信媒体通常包括计算机可读指令、数据结构、程序模块或经调制的数据信号（例如，载波或其他传送机制）中的其他数据，并且包括任何信息传送媒体。术语“经调制的数据信号”是指以对信号中的信息编码的方式来设置或改变其一个或多个特征。举例来讲（不作限制），通信媒体包括有线媒体（例如，有线网络或直接接线连接）和无线媒体（例如，声音、RF、红外线和其 他无线媒体）。上述媒体的任意结合也应包括在计算机可读媒体的范围中。

图 1 表现了一个这样的手持计算设备 120 的功能部件，包括处理器 122、存储器 124、显示器 126 和键盘 128（可以是物理键盘，也可以是虚拟键盘）。存储器 124 一般包括易失性存储器（例如，RAM）和非易失性存储器（例如，ROM、PCMCIA 卡等）两者。操作系统 130 驻留在存储器 124 中，并在处理器 122 上加以执行（例如，来自 Microsoft®公司的“Windows® CE”操作系统或另一种操作系统）。

一个或多个应用程序 132 被载入存储器 124 并在操作系统 130 上运行。应用程序的例子包括电子邮件程序、调度程序、PIM（个人信息管理）程序、文字处理程序、电子表格程序、互联网浏览器程序等。手持个人计算机 120 也可以包括加载在存储器 124 中的通知管理器 134，该通知管理器在处理器 122 上加以执行。通知管理器 134 处理例如来自应用程序 132 的通知请求。

手持个人计算机 120 具有由一节或多节电池构成的电源 136。电源 136 还可以包括使内置电池无效或再充电的外部电源（例如，AC 适配器或接通电源的连接底座）。

如图 1 中所示的示范手持个人计算机 120 示出三种类型的外部通知机制：一个或多个发光二极管（LEDs）140 和声频发生器 144。这些设备可以直接耦合到电源 136，以便当它们被激活时，即使关闭手持个人计算机处理器 122 和其他部件以保存电池功率，它们也可在通知机制所指示的期限内继续运行。较佳的是，LED 140 无限地继续保持，直到用户采取行动为止。注意，对于现今的手持个人计算机电池而言，声频发生器 144 的当代版本使用过多的功率，所以，当系统休眠时，或在激活之后的某段有限的时期，它被配置成关闭。

通用参考解码器

漏桶是作为时间函数的编码器或解码器缓冲器的状态（或充满度）的概念模型。图 2 表现了这个概念，其中，输入数据 200 被馈送到增强的编码器 202（在下文加以描述），该编码器把数据编码入编码器缓冲器 204。编码数据通过某个传输媒体（输送管）206 传输到解码器缓冲器 208，然后由解码器 210 解码成输出数据 212（例如，图像或视频帧）。为简单起见，这里将主要描述解码器缓冲器 208，因为编码器和解码器缓冲器的充满度从概念上讲是互补的，即，解码器缓冲器中的数据越多，编码器缓冲器中的数据就越少，反之亦然。

漏桶模型的特征是一组共三个参数—— R 、 B 和 F ，其中， R 是位进入解码器缓冲器 208 的峰值位速率（用位/秒表示）。在位速率恒定的情况下， R 经常是视频或声频片断的信道位速率和平均位速率，从概念上讲，可以认为它对应于输送管 206 的宽度。 B 是平稳视频位速率波动的桶或解码器缓冲器 208 的大小（用位表示）。这个缓冲器大小不能大于解码设备的物理缓冲器。 F 是初始的解码器缓冲器充满度（也用位表示），在解码器开始从缓冲器中除去位之前需要其存在。 F 至少与代表初始帧的编码数据的数量一样多。如果不考虑处理时间（对于本例来说，其目的在于可以认为它是瞬时的）， F 和 R 确定初始或启动延迟 D ，其中， $D=F/R$ 秒。

这样，在漏桶模型中，各个位以速率 R 进入解码器缓冲器 208，直到充满度水平为 F （即经 D 秒），然后（在本例中为立即）除去第一帧所需要的位 b_0 。

各个位以速率 R 持续进入缓冲器，解码器以某给定时间间隔除去后续帧的 b_1 、 $b_2 \dots b_{n-1}$ 位，该间隔通常（但不一定）是每 $1/M$ 秒，其中的 M 是视频的帧速率。

图 3 是说明位流的解码器缓冲器充满度与时间的图表，该位流包含在漏桶的参数 (R, B, F) 中，如上所述，其中第 i 个帧的位的数量是 b_i 。在图 3 中，从缓冲器中除去经编码的视频帧（通常根据视频帧速率），如缓冲器充满度降低所示。具体说，在时间 t_i 除去 b_i 位之前，立刻让 B_i 成为解码器缓冲器充满度。一般的漏桶模型根据以下的公式来进行操作：

$$B_0 = F$$

$$B_{i+1} = \min (B, B_i - b_i + R(t_{i+1} - t_i)), \quad i=0, 1, 2, \dots$$

通常， $t_{i+1} - t_i = 1/M$ 秒，其中的 M 是位流的帧速率（用帧/秒表示）。

如果解码器缓冲器 208 不发生下溢（图 2），则具有参数 (R, B, F) 的漏桶模型包含一个位流。由于编码器与解码器缓冲器充满度彼此互补，因此，这等同于编码器缓冲器 204 没有溢出。但是，允许编码器缓冲器 204（漏桶）变空，或解码器缓冲器 208 同样可以变满，这时，就没有位进一步从编码器缓冲器 204 被传输到解码器缓冲器 208。这样，当解码器缓冲器 208 充满时，它就停止接收位，这就是在以上第二个方程式公式中使用减运算符的原因。如以下关于可变的位速率（VBR）流的描述所述，由于它们是互补的，因此，充满的解码器缓冲器 208 意味着编码器缓冲器 204 是空的。

注意，一个给定的视频流可能包含在各种漏桶配置中。例如，如果视频流包含在具有参数 (R, B, F) 的漏桶中，那么，它也将被包含在具有一个更大的缓冲器 (R, B', F) 的漏桶中，其中， B' 比 B 大；或者，它将被包含在具有更高的峰值传输速率 (R', B, F) 的漏桶中，其中， R' 比 R 大。此外，对于任何位速率 R' ，存在一个缓冲器大小将包含（时限）视频位流。在最差的情况下，即当 R' 接近零时，缓冲器大小将需要和位流本身一样大。换言之，只要缓冲器大小足够大，就可以按任何速率（不管片断的平均位速率如何）来传输视频位流。

图 4 是关于一个给定的位流的最小缓冲器大小 B_{\min} 与峰值位速率 R_{\min} 的比较图表，这采用了以上第二个公式，其中，将所需的初始缓冲器充满度设定在总缓冲器大小的一个恒定部分。图 4 中的曲线指出：为了以峰值位速率 r 传输流，解码器需要缓冲至少 $B_{\min}(r)$ 个位。此外，从图表中可以理解，更高的峰值速率要求较小的缓冲器大小，因此，缩短了启动缓冲器延迟。或者，

图表指出：如果解码器缓冲器的大小是 b ，则传输位流所要求的最小的峰值速率是有关的 $R_{\min}(b)$ 。而且，任何位流（例如，图 4 中的位流）的 (R_{\min}, B_{\min}) 对的曲线呈分段线性和中凸。

根据本发明的一个方面，如果增强的编码器 202 提供至少两个曲线点，则通用参考编码器 210 可以选择一个点，或线性地在各个点之间内插，或外插各个点，以达到略（但安全地）大于 (R_{\min}, B_{\min}) 的一些点 $(R_{\text{interp}}, B_{\text{interp}})$ 。一个重要的结果是，在许多情况下，可以安全地减小缓冲器大小约相对于单个漏桶的一个数量级，该漏桶含有按其平均速率的位流，由此，同样减少了延迟。或者，对于相同的延迟，峰值传输速率可以减小（可能）4 倍，或信号-噪声比（SNR）可提高（可能）几个 dB。

为此，通过安排编码器 202 生成至少两组漏桶参数 214（例如， (R_1, B_1, F_1) 、 (R_2, B_2, F_2) 、... (R_N, B_N, F_N) ），来增强编码器 202，这些漏桶参数 214 对应于速率-缓冲器曲线上的至少两个点，它们对于给定的视频或图像片断很有用（例如，相对于 R 和/或 B 的范围方面可合理地分开）。然后，例如通过将这些漏桶参数设置内插入初始流头部中或利用某种频带以外的方式，增强的编码器 202 将它们和其数量 N 一起提供给通用参考解码器。注意，即使对于相对大的 N （例如，几十个桶，而 2-4 个桶通常将足以合理地代表 R - B 曲线）而言，当与典型的视频或图像数据进行比较时，也可以忽略提供该信息所必要的额外字节的数量（例如， N 的一个字节，加上 8 个字节/漏桶模型或参数设置）。

此外，注意，在较高的位速率下，内容创建者可以决定在位流中的不同时间规定不同的漏桶模型，这在无论何时连接在传输期间失败并在位流中间重新开始时都很有用。例如，可以在 15 分钟的间隔提供漏桶模型，这样，通过重新选择、在合适的时间重新内插或重新外插，解码器可以按需要来改变其操作条件（例如，它的缓冲器大小或速率）。

可以由编码器来选择所需的值 N （注意，如果 $N=1$ ，则通用解码器 210 将外插类似 MPEG 视频缓冲检验符的点）。编码器可以选择预先选择各个漏桶值并利用速率控制（确保符合各种漏桶限制）对位流进行编码，先对位流编码，然后使用上述公式来计算各组漏桶参数（包含处于 R 的 N 个不同值的位流），或进行这两种操作。第一种方法可以被应用于实况或即时（on-demand）传输，其他的方法可应用于即时传输。

在执行本发明的过程中，一旦在通用参考解码器 210 处被接收，如果知道它可用的峰值位速率和/或它的物理缓冲器大小，解码器 210 就可以确定它希望使用哪个漏桶。或者，通用参考解码器 210 可以线性地在其间内插或线性地从这些点进行外插，以便为给定的配置找出一组合适的参数。图 5 表现了两种漏桶参数组及其线性内插 (R, B) 值。为参考起见，经计算的 R-B 曲线用细的虚线表示，而漏桶模型 (R_x, B_x) 和 (R_y, B_y) 中所提供的 R 和 B 值则用星号表示。从 (R_x, B_x) 到 (R_y, B_y) 的实线代表被插入的值。实线上所选的任何 R 或 B 配对将适当地保持（例如，不溢出或下溢）解码器缓冲器 208。也可以从这些点外插漏桶参数，（由图 5 中粗的虚线表示）并且，实线上所选的任何 R 或 B 配对还将适当地保持解码器缓冲器 208。

在点 k 与 k+1 之间内插的缓冲器大小 B 遵从直线：

$$B = \frac{R - R_k}{R_{k+1} - R_k} B_k + \frac{R_{k+1} - R}{R_{k+1} - R_k} B_{k+1}$$

其中， $R_k < R < R_{k+1}$ 。

同样，可以线性地插入初始的解码器缓冲器充满度 F：

$$F = \frac{R - R_k}{R_{k+1} - R_k} F_k + \frac{R_{k+1} - R}{R_{k+1} - R_k} F_{k+1}$$

其中， $R_k < R < R_{k+1}$ 。

保证所产生的具有参数 (R, B, F) 的漏桶包含位流，因为（用数学方法可证明）最小的缓冲器大小 B_{min} 在 R 和 F 中是中凸的，也就是说，对应于任何中凸组合 $(R, F) = a(R_k, F_k) + (1 - a)(R_{k+1}, F_{k+1})$ （其中， $0 < a < 1$ ）的最小的缓冲器大小 B_{min} 小于或等于 $B = aB_k + (1 - a)B_{k+1}$ 。

如上所述，当 R 大于 R_N 时，漏桶 (R, B_N, F_N) 也将包含位流，从而当 $R \geq R_N$ 时，B_N 和 F_N 是所推荐的缓冲器大小和初始的解码器缓冲器充满度。如果 R 小于 R₁，则可以使用上限 $B = B_1 + (R_1 - R)T$ ，其中 T 是用秒表示的流的时间长度。可以外插在 N 个点的范围以外的这些 (R, B) 值。

应该注意，解码器不需要选择、内插或外插漏桶参数，但另一个实体可以选择用于将单个的设置发送到解码器的各种参数，然后，它将使用那一个设置。例如，通过提供诸如解码器的要求等信息，服务器可以确定（通过选择、内插或外插）一组合适的漏桶参数，以发送给解码器，然后，解码器可以只使用单个的一组参数来进行解码。服务器或解码器的代理也能够选择、

内插或外插漏桶信息，而解码器无须查看一个以上的漏桶。换言之，服务器可以代替解码器作决定，服务器和客户解码器可能会对各个参数进行商议。但是，一般而言，与本发明一致，根据至少两个漏桶模型来事先或动态地确定一个合适的漏桶模型。

可以从最高和最低的充满度值在解码器缓冲器曲线中出现的时间，来计算关于给定位流的 R-B 曲线的各个值（例如如图 3 中所示的那些值）。具体说，对一个漏桶的参数（R, B, F）漏桶中所包含的位流分别考虑解码器缓冲器充满度的最高值和最低值的两个时间（ t_M , t_m ）。可以在几种场合中达到充满度的最高值和最低值，但要考虑 $t_M < t_m$ 时的最大值（ t_M , t_m ）。假设适当地计算漏桶，B 是包含值 R、F 的位流的最小的缓冲器大小，那么，

$$B = \sum_{t=t_M}^{t=t_m} \left[b(t) - \frac{R}{M} \right] = \sum_{t=t_M}^{t=t_m} b(t) - n \frac{R}{M} = -n \frac{R}{M} + c, \quad ,$$

其中， $b(t)$ 是在时间 t 时的帧的位的数量， M 是用帧/秒表示的帧速率。在这个公式中， n 是时间 t_M 与 t_m 之间的帧的数量， c 是那些帧的位的总数。

这个公式可以作为一个点被插入直线 $B(r)$ 中，其中， $r = R$ ，且 $-n/M$ 是该线的斜度。有一个位速率范围 $r \in [R-r_1, R+r_2]$ ，以便最大的一对值 t_M 和 t_m 将保持相同，从而以上的公式对应于定义与位速率 r 有关的最小缓冲器大小 B 的一条直线。如果位速率 r 在以上的范围以外，则值 t_M 和/或 t_m 中的至少一个值将发生变化，由此，如果 $r > R+r_2$ ， t_M 与 t_m 之间的时间间隔将更小，定义 $B(r)$ 的新直线中的值 n 也将更小，各条线的斜度将更大（较小负值）。如果 $r < R-r_1$ ， t_M 与 t_m 之间的时间间隔将更大，定义 $B(r)$ 的直线中的值 n 也将更大。于是，该线的斜度将更小（更大负值）。

位速率范围的各个对（ t_M , t_m ）的值（或有关 n 的值）和 c 的一些值（对于给定的一对至少有一个值）可以存储在位流的头部中，这样，可以使用以上公式来获得分段线性 $B(r)$ 曲线。此外，在编码器已经生成位流之后，这个公式可以用来简化漏桶模型参数的计算。

在测试中，产生图 5 中的位流，形成 797 Kbps 的平均位速率。如一般在图 5 中所示，在 797 Kbps 的恒定传输速率下，解码器将需要约 18,000 Kbits (R_x , B_x) 的缓冲器大小，初始的解码器缓冲器充满度等于 18,000 Kbits，启

动延迟大约为 22.5 秒。这样，该编码（不利用速率控制器而产生）变换位直到 22.5 秒，以便对其整体的编码长度本质上实现最佳的、可能的质量。

图 5 也表现了在 2,500 Kbps 的峰值传输速率（例如，2x CD 的视频位速率部分）下，解码器将只需要 2,272 Kbits 的缓冲器大小（ R_y , B_y ），这对于消费级硬件设备而言是合理的。当初始的缓冲器充满度等于 2,272 Kbits 时，启动延迟将只有大约 0.9 秒。

这样，对于这个编码而言，两个漏桶模型通常可能是有用的，例如，（ $R=797$ Kbps, $B=18,000$ Kbits, $F=18,000$ Kbits）和（ $R=2,500$ Kbps, $B=2,727$ Kbits, $F=2,272$ Kbits）。这第一个漏桶参数组将允许视频在一个恒定的位速率信道上传输，延迟大约为 22.5 秒。而这个延迟对于许多情况而言可能太大，但它对于例如互联网的电影流而言也许可以接受。第二组漏桶参数将允许视频在具有峰值速率为 2,500 Kbps 的共享网络上传输，或者将允许从 2x CD 进行本地重放（延迟约为 0.9 秒）。这种 1 秒之下的延迟对于通过类似 VCR（盒式磁带录象机）的功能性的随机存取重放而言是可接受的。

当考虑到在位流中只规定第一个漏桶而不是第二个漏桶时会发生什么情况，就会明白各种好处。在这种情况下，即使当在具有峰值位速率 2,500 Kbps 的信道上进行重放时，解码器也将使用大小为 18,000 Kbits 的缓冲器，这样，延迟将为 $F/R = 18,000 \text{ Kbits} / 2,500 \text{ Kbps} = 7.2$ 秒。可以理解，对于随机存取重放（例如，具有类似 VCR 的功能性）而言，这种延迟是不可接受的。但是，如上所述，如果附加地规定第二个漏桶，那么，在速率为 2,500 Kbps 时，缓冲器大小下降到 2,272 Kbits，延迟下降到 0.9 秒。

另一方面，如果只规定第二个漏桶（而不是第一个漏桶），那么，在恒定的传输速率 797 Kbps 下，即使是智能解码器也会被迫使用远远大于必要大小的缓冲器，以确保缓冲器不会溢出，即， $B' = B + (R - R') T = 2,272 \text{ Kbits} + (2,500 \text{ Kbps} - 797 \text{ Kbps}) \times 130 \text{ 秒} = 223,662 \text{ Kbits}$ 。即使在给定设备中可获得如此多的存储器，但这对应于 282 秒的初始延迟或接近 5 分钟，这是无法接受的。但是，如上所述，如果也规定了第一个漏桶，那么，在 797 Kbps 的速率下，缓冲器大小下降到 18,000 Kbits，延迟下降到 22.5 秒。

而且，当规定两个漏桶时，对 797 Kbps 到 2,500 Kbps 之间的任何位速率，解码器可以在两个漏桶之间进行线性插入（使用以上的插入公式），从而在任何给定的速率都实现近似最小的缓冲器大小和延迟。与在 797 Kbps 到

2,500 Kbps 之间只具有单个的漏桶的外插相比, 外插(在图 5 中由粗虚线表示)在 797 Kbps 以下和 2,500 Kbps 以上也更加有效率。

如上述例子所示, 甚至仅有两组漏桶参数就可以提供缓冲器大小方面数量级的减小(例如, 在一种情况下, 从 223,662 到 18,000 Kbits; 在另一种情况下, 从 18,000 到 2,272 Kbits), 并在给定的峰值传输速率时, 可以提供使延迟有数量级的减少(例如, 在一种情况下, 从 281 到 22.5 秒; 在另一种情况下, 从 7.2 到 0.9 秒)。

或者, 也可以对给定的解码器缓冲器大小减小峰值传输速率。实际上, 从图 5 中显而易见, 如果可以通过内插和/或外插多个漏桶来获得 R-B 曲线, 那么, 具有固定的物理缓冲器大小的解码器就可能选择能安全地为位流解码所需要的最小峰值传输速率, 而不会发生解码器缓冲器下溢。例如, 如果解码器有一个大小为 18,000 Kbits 的固定的缓冲器, 则用于编码的峰值传输速率可以是低至 797 Kbps。但是, 如果只规定第二个漏桶(而不是第一个漏桶), 那么, 解码器可以将位速率减小到不小于 $R' = R - (B' - B) / T = 2,500 \text{ Kbps} - (18,000 \text{ Kbits} - 2,272 \text{ Kbits}) / 130 \text{ 秒} = 2,379 \text{ Kbps}$ 。在这种情况下, 对于相同的解码器缓冲器大小而言, 与使用单个的漏桶相比, 只使用两个漏桶可将峰值传输速率减小 4 倍。

在相同的平均编码速率下, 具有多个漏桶参数也可以改善重建的视频的质量。考虑编码过程可用两个漏桶的情况。如上所述, 利用解码器处的该信息, 就可以重放编码过程, 如果峰值传输速率是 797 Kbps, 则延迟是 22.5 秒; 如果峰值传输速率是 2,500 Kbps, 则延迟是 0.9 秒。

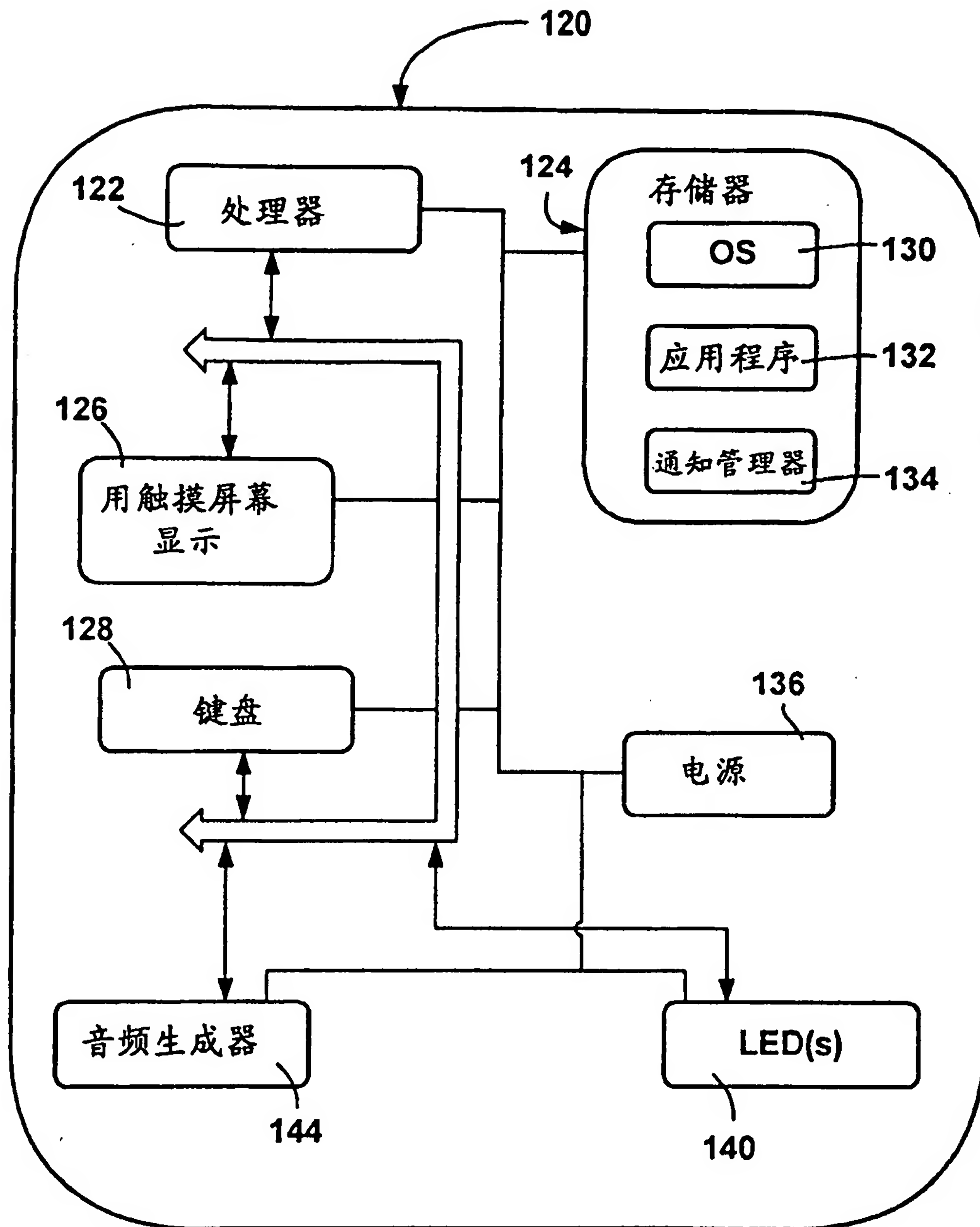
但是, 如果不具备第二个漏桶, 则在 2,500 Kbps 的条件下, 延迟从 0.9 秒增加到 7.2 秒。在没有第二个漏桶的条件下将延迟减回到 0.9 秒的一种方法是: 通过将(第一个漏桶的)缓冲器大小从 18,000 Kbits 减少到 $(0.9 \text{ 秒}) \times (2,500 \text{ Kbps}) = 2,250 \text{ Kbits}$, 来利用速率控制对片断重新编码。这将确保: 如果峰值传输速率是 2,500 Kbps, 则延迟只有 0.9 秒, 但 797 Kbps 处的延迟也将从 22.5 秒减少到 2.8 秒。但是, 结果是, 质量(SNR)也将下降一定量——估计是几个 dB, 尤其是对具有大的动态范围的片断而言。

这样, 规定第二个漏桶可以将 SNR 增加(可能)几个 dB, 除了规定第二个漏桶的可忽略数量的额外的位/片断以外, 平均位速率不会发生变化。在每个峰值传输速率的重放上可察觉这个 SNR 的增加。

当单个的编码在具有不同峰值速率的信道上加以传输或传输到具有不同的物理缓冲器大小的设备时，为广义通用参考解码器规定多个漏桶的好处便实现了。但是，在实际应用中，这正在变得越来越普通。例如，被进行离线编码并被存储在磁盘上的内容经常被进行本地重放，也经常在具有不同峰值速率的网络上传输。即使对于本地重放来说，不同的驱动器速度（例如，1xCD~8xDVD）也会影响峰值传送速率。而且，通过网络连接的峰值传输速率也根据限制链路的速度而动态地发生变化，限制链路通常在最终用户附近（例如，100 或 10 baseT 以太网、T1、DSL、ISDN、调制解调器等）。重放设备的缓冲器容量也有显著的变化——从具有千兆字节的缓冲器空间的桌上型计算机，到具有小几个数量级的缓冲器空间的小型消费电子设备。利用本发明的多个漏桶和所提议的通用参考解码器，相同的位流可以在各种信道上以最少启动延迟，最小解码器缓冲器要求，最大可能的质量进行传输。这不仅可应用于离线编码的视频，而且可应用于通过不同的信道同时向不同的设备广播的实况视频。简而言之，所提议的广义通用参考解码器为现存的位流增添了显著的灵活性。

从前面详细的描述中可见，与原先标准中的解码器相比，提供了一种改进的通用参考解码器。该广义通用参考解码器仅从编码器要求少量的信息（例如，在位流的头部），以便为位流通过现今网络（其中，带宽是可变带宽）和/或终端（具有各种位速率和缓冲容量）的传输提供了更高的灵活性。本发明的参考解码器实现了这些新的方案，而为使对可用的带宽的传输延迟减少到最小值，此外，实际上将传送到设备（具有给定的物理缓冲器大小限制）的信道位速率要求减到最少。

本发明容许各种修改和替代构造，其某些所展示的实施例在附图中被示出并已经在上文详细地描述过。但是，应该理解，这不是意在将本发明局限于所揭示的各种具体的形式，而相反，本发明包括其精神和范围内的所有修改、替代构造和同等物。



图

1

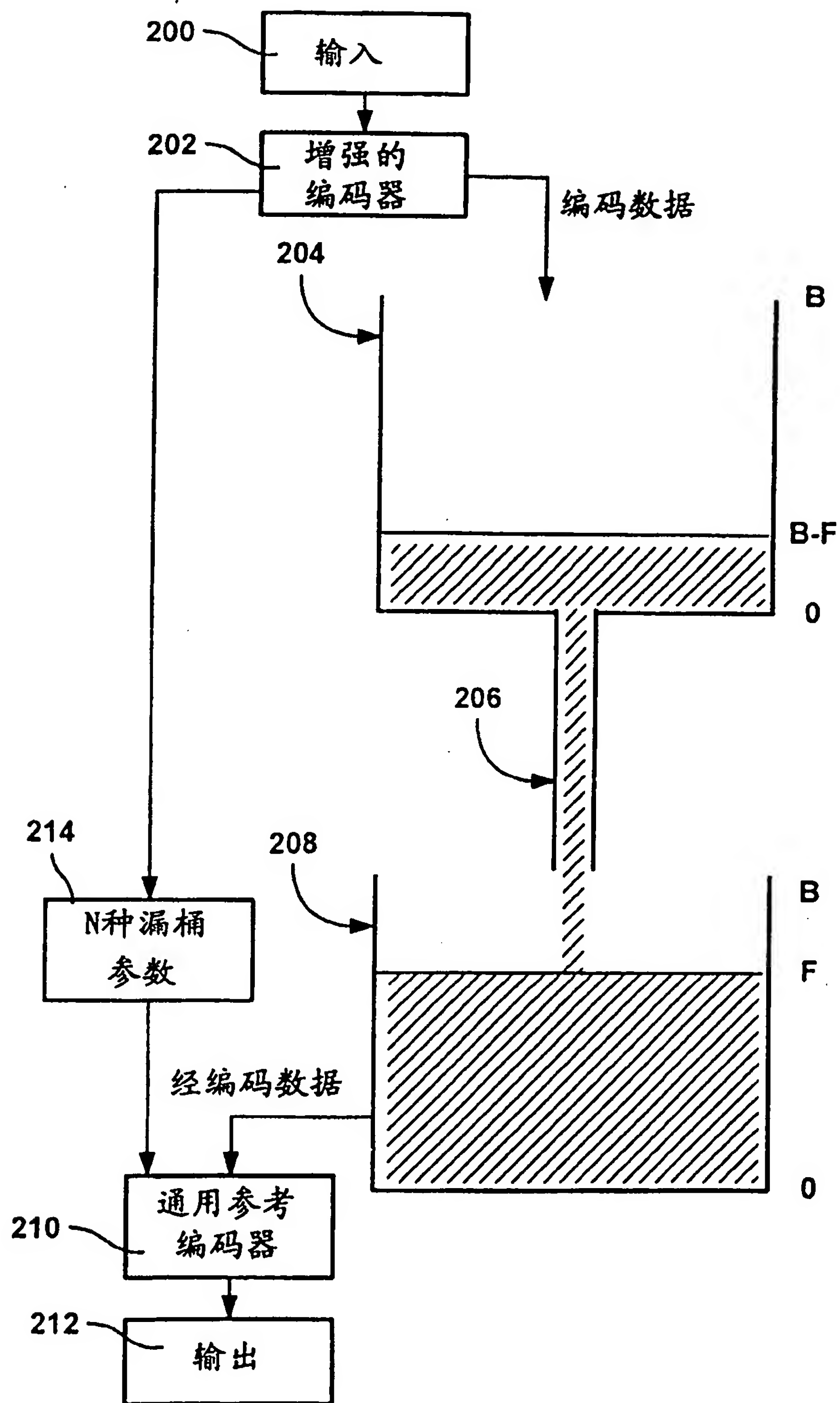


图 2

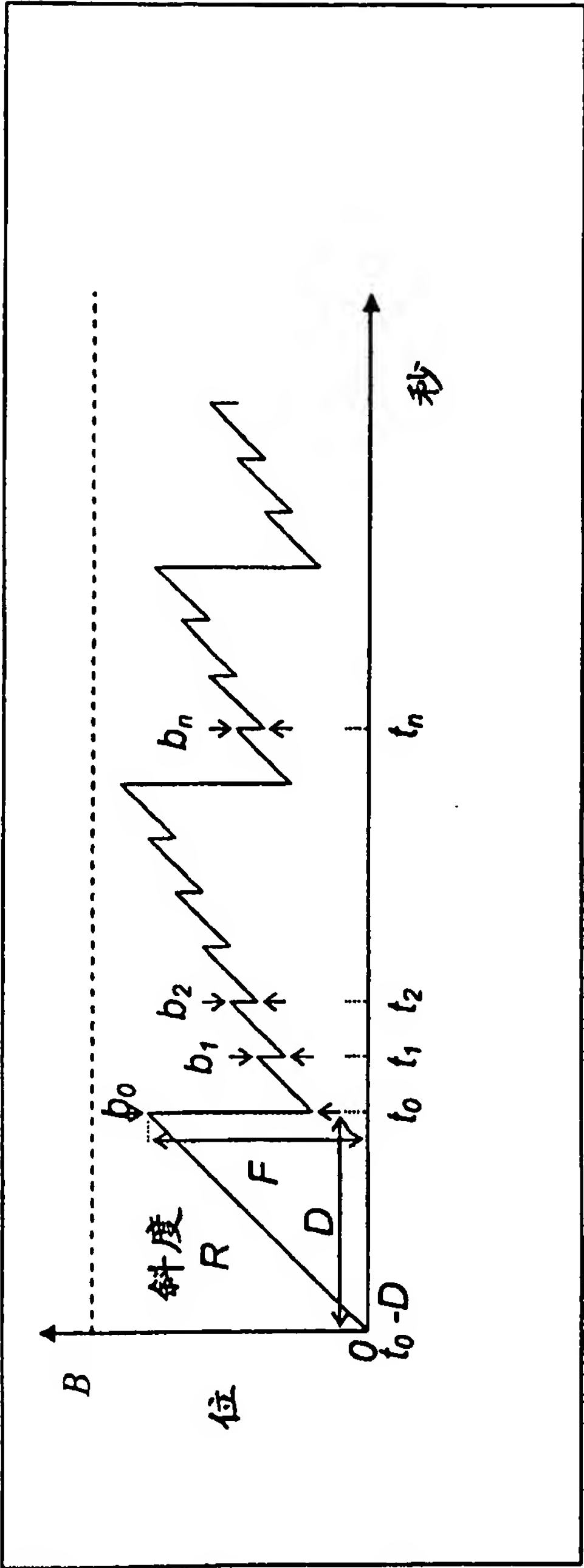


图 3

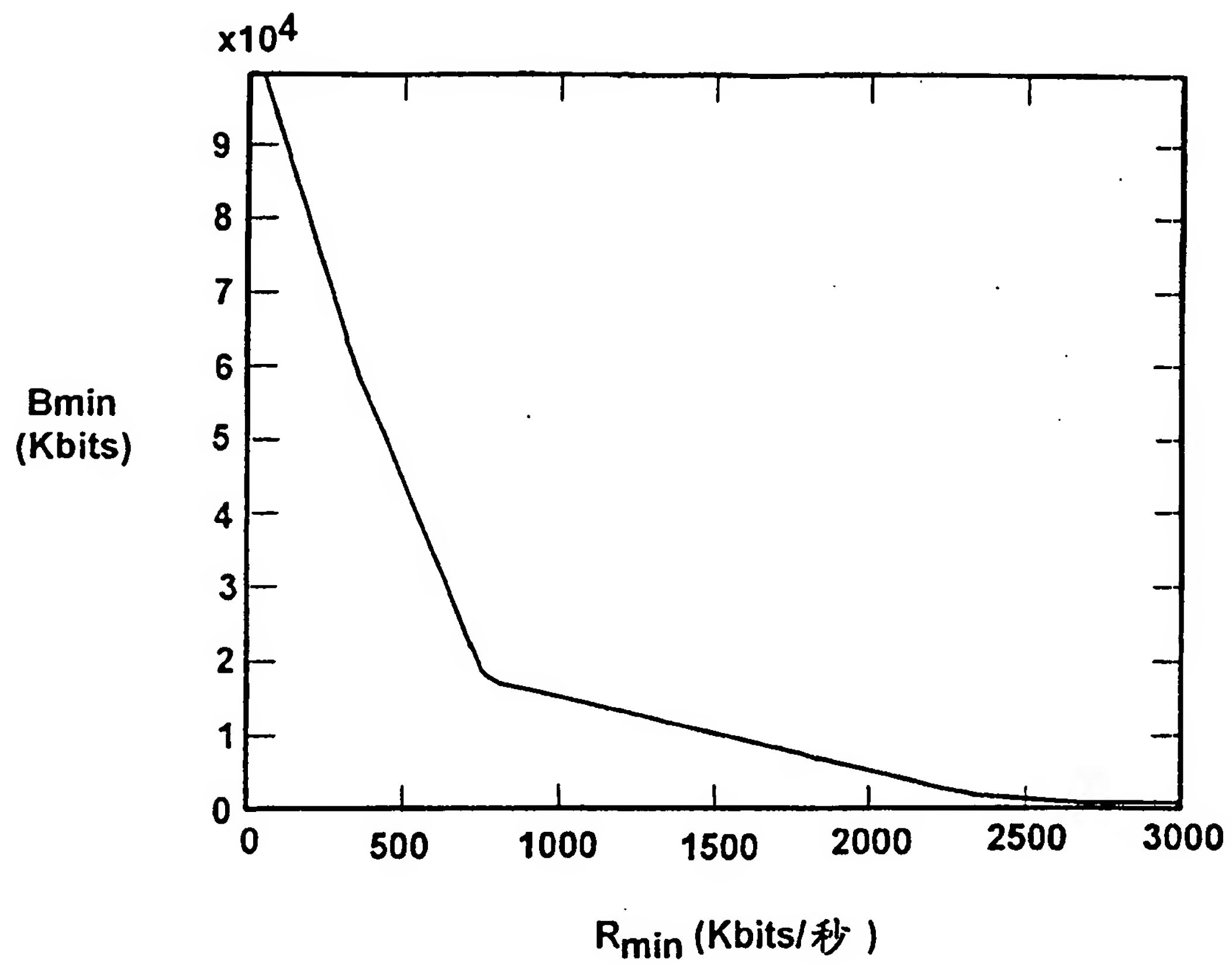


图 4

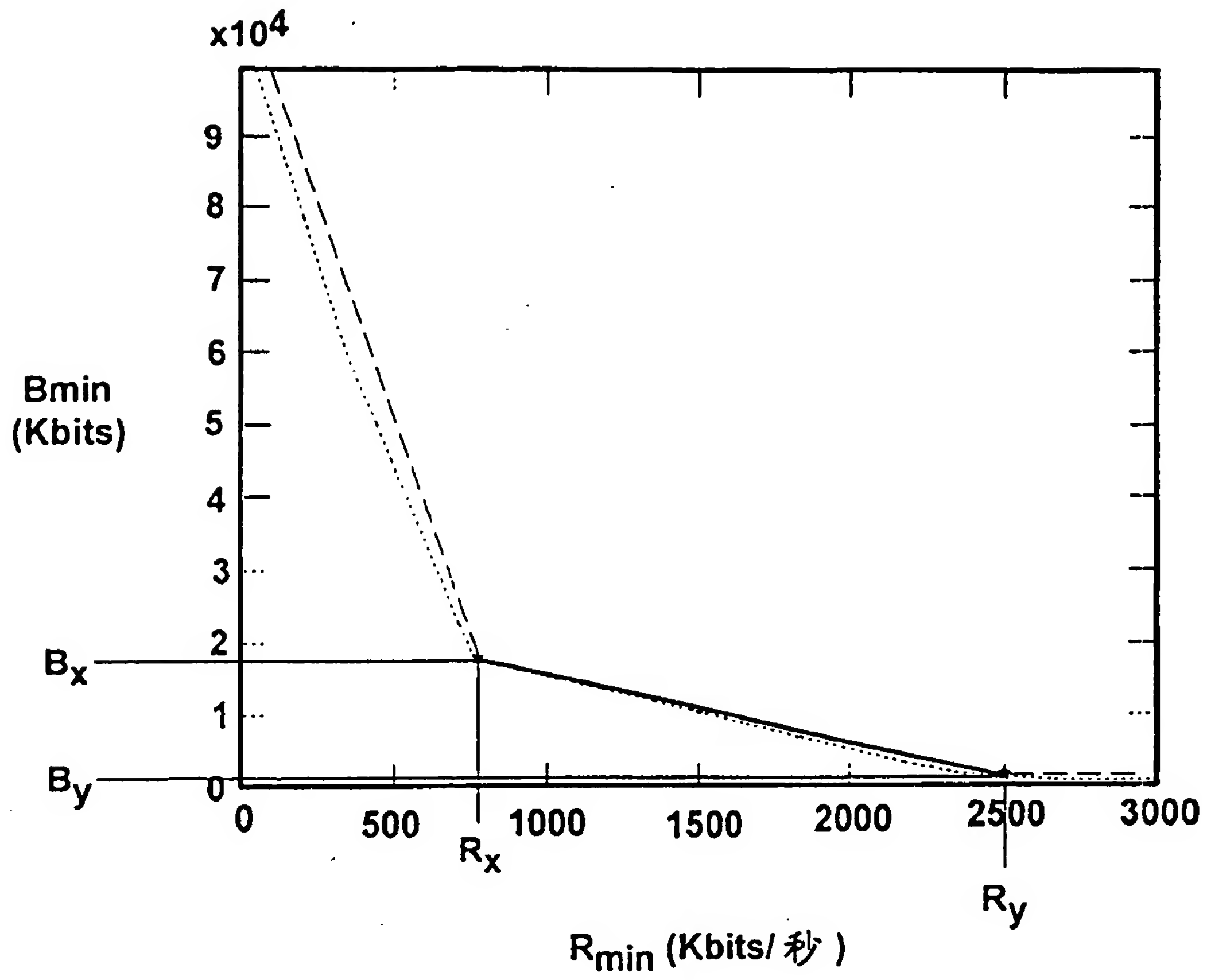


图 5